causes: (*a*) If there are N numbers, there should be found $f(x)N$ of which the absolute deviation is equal to or greater than x; theory gives the value of $f(x)$; (*b*) the value of the expression $\frac{2N\Sigma\delta^2}{(\Sigma\delta)^2}$ should be 3.14159 . . . Now meteorological data may satisfy both these tests without at all fulfilling other conditions equally demanded by theory; we have here a good illustration of the oft-repeated warning against drawing conclusions from summary coefficients alone, such as the mean. In the present instance, the *order* in which the numbers appear is of great significance, and the following relation must also hold:[4]

If the deviations from the mean are to be likened to fortuitous errors, then the ratio of the mean variability to the mean deviation must be equal[5] *to* $\sqrt{2}=1.414$ . . . The variabilities and deviations are taken without regard to sign.

Drawings from a sack containing balls, on each of which was marked an observed daily temperature, would give a succession vastly different from the succession actually observed: Long series of increasing or decreasing values would be less frequent in the drawing than in the observing, and the mean variability would be greater in the former; in fact the ratio of mean variability to mean deviation in the case of series of daily temperatures turns out to be but little more than half the theoretical value; chance would give the deviations which are observed, but would not give the succession which is observed. *Yet both the actual and the chance successions satisfy the two tests mentioned above.*

It has been pointed out by Besson (*op. cit.*) that if a variable is taking on random values, it does not follow that the succession of the signs of the variations will obey the laws of chance; Goutereau points out further that the deviations from the mean may not be fortuitous even if they follow the Law of Gauss.—*Edgar W. Woolard.*

[4] Ch. Goutereau: Sur la variabilité de la température, *Annuaire de la Soc. Mét. de France*, 54, 122–127, 1906.
[5] The demonstration, by Maillet, is given by Goutereau, *op. cit.* The absolute difference between a number and the next consecutive number is the variability.

551.54 : 551.501

THE VARIATE-DIFFERENCE CORRELATION METHOD.

For correlating daily changes of barometric height at Halifax and Wilmington, Miss Cave[1] made use of a formula, devised by Pearson, giving the correlation coefficient between the differences of successive daily readings at the two stations; and remarked that this formula would apply to any case in which it was desired to correlate the difference of one pair of quantities with the difference of another pair; no comments on where this procedure might be desirable were offered, however. Later, Hooker[2] independently pointed out that the correlation coefficient between two variables, for each of which a series of observations is available, is a test of similarity of the two phenomena as influenced by the totality of the causes affecting each of them; when, therefore, the observations extend over a considerable period of time, certain difficulties arise which find no precise parallel in the case where the whole of the observations refer to the same moment of time: If a diagram be drawn, showing by curves the changes of the two variables during the period under consideration, some relation will often be suggested between the usually smaller and more rapid alterations while at the same time the slower "secular" changes may or may not exhibit any similarity. If, then, the correlation coefficient be formed in the ordinary way, employing deviations from the mean, a high value will be obtained if the "secular" changes are similar (this value being almost independent of the similarity or dissimilarity of the more rapid changes), but a value approximating to zero if the "secular" changes are of quite dissimilar character even though the similarity of the smaller rapid changes be extremely marked; deductions drawn from ordinary correlation coefficients may be very erroneous. In order to get rid of the spurious correlation arising from the fact that both variables are functions of the time, the correlation coefficient may be formed between the *variations*, or first differences, of the quantities, instead of between the quantities themselves. After this method had been in rather extensive use for some time, Pearson pointed out that it was valid only when the connection between the variables and the time was linear.

The name Variate-Difference Correlation was given by Pearson[3] to a generalization of the preceding artifice, in which it was demonstrated[4] that if the variables are randomly distributed in time and space, the correlation between the variables and that between the corresponding *nth* differences will be the same; and that when this is not the case, we can eliminate variability which is due to position in time or space, and so determine whether there really is any correlation between the variables themselves, by correlating the 1st, 2d, 3d, * * *, *nth* differences: *when the correlations between the differences remain steady for several successive orders of differences we may reasonably suppose we have reached the true correlation between the variables.*

The complete theory of the method was worked out by Anderson[5] and subjected to critical examination by Pearson (*op. cit.*), who found that, as usual, the theoretical formulæ were only roughly approximated to in practice unless a great number of observations were at hand.

There has been no source more fruitful of fallacious statistical argument than the common influence of the time factor. The difference method of correlation is one of great promise and usefulness. The very frequent and superficial statements that such and such variables, both changing rapidly with the time, are essentially causative cease to have any foundation when the difference method is applied.[6]—*Edgar W. Woolard.*

[1] F. E. Cave-Browne-Cave; On the influence of the time factor on the correlation between the barometric heights at stations more than 1,000 miles apart, *Proc. Roy. Soc.*, 74:403–413, 1904–1905.
[2] R. H. Hooker: On the correlations of successive observations, *Jour. Roy. Statistical Society*, 68:696–703, 1905.

[3] Beatrice M. Cave and Karl Pearson: Numerical illustrations of the variate difference correlation method, *Biometrika*, 10, 340–355, 1914–15.
[4] "Student": The elimination of spurious correlation due to position in time or space, *Biometrika*, 10, 179–180, 1914–15.
[5] Nochmals über "The elimination of spurious correlation due to position in time or space," O. Anderson, *Biometrika*, 10, 269–279, 1914–15.
[6] Illustrations of the method are given by Cave and Pearson, *op. cit.*, and by G. U. Yule, *Introduction to the Theory of Statistics*, 5 ed., 1919, pp. 197–201; see also T. Okada, Some researches in the far eastern seasonal correlations, MO. WEATHER REV., 1917, 45: 238, 299, 535.

NOTE ON PROF. MARVIN'S DISCUSSION OF "A POSSIBLE RAINFALL PERIOD EQUAL TO ONE-NINTH THE SUN-SPOT PERIOD."

By DINSMORE ALTER.

[University of Kansas, Lawrence, Kans., Apr. 26, 1921.]

I have naturally been much interested in Prof. Marvin's conclusions[1] regarding my paper.[2] I am very sorry that it is impossible for us to agree concerning the possibility of the phenomenon discussed, and especially concerning the legitimacy of the method employed. A further statement concerning some of the points raised by him may be in order.

[1] MO. WEATHER REV., February, 1921, 49: 83–85.
[2] *Ibid.*, pp. 74–83.